Average-case performance of heuristics for three-dimensional random assignment problems

Alan Frieze* Gregory B. Sorkin†

2 Aug 2010

Abstract

Beautiful formulas are known for the expected cost of random two-dimensional assignment problems, but in higher dimensions, even the scaling is not known. In 3 dimensions and above, the problem has natural "planar" and "axial" versions, both of which are NP-hard. For 3-dimensional Planar random assignment instances of size n, the cost scales as $\Omega(1/n)$, and a main result of the present paper is the first polynomial-time algorithm that, with high probability, finds a solution of cost $O(n^{-1+\varepsilon})$, for arbitrary positive ε (or indeed ε going slowly to 0). For 3-dimensional Axial assignment, the lower bound is $\Omega(n)$, and we give a new efficient matching-based algorithm that returns a solution with expected cost $O(n \log n)$.

1 Introduction

An instance of the (two dimensional) assignment problem may be thought of as an $n \times n$ cost array $C_{i,j}$, a candidate solution is a permutation $\pi\colon [n] \mapsto [n]$, its cost is $\sum_{i=1}^{n} C_{i,\pi(i)}$, and an optimal solution is one minimizing the cost. If the cost matrix represents, for example, the costs of assigning various jobs i to machines j, where each machine can accommodate only one job, then the problem's solution represents the cheapest way of assigning the jobs to machines. It may equivalently be formulated as an integer linear program, minimizing the sum of selected elements consistent with the selection of exactly one element from each row and from each column, i.e., minimizing $\sum_{i,j} X_{i,j} C_{i,j}$ where $X_{i,j} \in \{0,1\}$, $(\forall i)$ $\sum_j X_{i,j} = 1$ and $(\forall j)$ $\sum_j X_{i,j} = 1$. This is a network flow problem, thus its linear relaxation with $X_{i,j} \in [0,1]$ has integer extreme points, and the problem may be solved in polynomial time.

The *random assignment problem*, in its most popular form, is the case when the entries of the cost matrix C are i.i.d. Exp(1) random variables (independent, identically distributed exponential random variables with parameter 1). Since anyway the problem can be solved in polynomial time, the focus for the random case is on the cost's expectation as a function of n,

$$f(n) = \mathbf{E}\left[\min_{\pi} \sum_{i=1}^{n} C_{i,\pi(i)}\right] = \mathbf{E}\left[\min_{X_{i,j}} \sum_{i,j} X_{i,j} C_{i,j}\right]$$

with $X_{i,j} \in \{0,1\}$ subject to the row and column constraints as above. This problem has received a great deal of study over several decades. It was considered from an operations research perspective in the 1960s [Don69], an asymptotic conjecture $f(n) \to \pi^2/6 = \zeta(2)$ was formulated by statistical physicists Mézard and Parisi in the 1980s based on the mathematically sophisticated but non-rigorous "replica method" [MP85, MP87], an exact conjecture $f(n) = \sum_{i=1}^{n} 1/i^2$ was hazarded by

*Research supported by NSF grant DMS-6721878, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, e-mail alan@random.math.cmu.edu

†Department of Mathematical Sciences, IBM T.J. Watson Research Center, Yorktown Heights NY 10598, e-mail sorkin@watson.ibm.com

Parisi in the late 1990s [Par98] and a generalization to partial matchings and non-square matrices made by Coppersmith and Sorkin [CS99], the Mézard–Parisi conjecture was proved by Aldous in a pair of papers in 1992 and 2001 [Ald92, Ald01], the Coppersmith–Sorkin conjecture was proved simultaneously in 2004 by two papers using two very different methods [NPS05, LW04], and the study of other aspects of the random assignment problem and related problems is ongoing, for example by Wästlund in [Wäs09].

In higher dimensions there are two natural generalizations of the assignment problem. For example in three dimensions, the Planar assignment problem is, given an $n \times n \times n$ matrix (or "tensor" or "array") C, to find a solution $X_{i,j,k}$ minimizing $\sum_{i,j,k} X_{i,j,k} C_{i,j,k}$ where $X_{i,j,k} \in \{0,1\}$ and there is one selected value per "plane" of the array:

$$(\forall i) \sum_{j,k} X_{i,j,k} = 1, \quad (\forall j) \sum_{i,k} X_{i,j,k} = 1, \quad (\forall k) \sum_{i,k} X_{i,j,k} = 1.$$

Equivalently, it is to determine $\min_{\pi,\sigma} \sum_{i=1}^{n} C_{i,\pi(i),\sigma(i)}$, the minimum taken over a pair of permutations. The Axial three dimensional assignment problem is similar but with one selected value per "line" of the array:

$$(\forall i,j) \sum_{k} X_{i,j,k} = 1, \quad (\forall j,k) \sum_{i} X_{i,j,k} = 1, \quad (\forall i,k) \sum_{j} X_{i,j,k} = 1.$$

The generalizations to higher dimensions are clear. In 3 dimensions and higher, the Planar and Axial assignment problems are both NP-hard. The Planar case $d = 3$ was one of the original problems listed by Karp [Kar72]. The complexity of the Axial problem was established by Frieze [Fri83].

The *multidimensional random assignment problem* we consider here is the case when the entries of the cost matrix are i.i.d. Exp(1)random variables. In this random setting, there are two natural questions. First, are there polynomial-time algorithms that find optimal or near-optimal solutions **whp**? Second, what is the expected cost of a minimum assignment? A random two-dimensional assignment instance has limiting expected cost $\zeta(2)$, and Frieze showed that the expected cost of a minimum spanning tree in the complete graph with random Exp(1) edge weights tends to $\zeta(3)$ [Fri85], so it is tantalizing to wonder if there might be similarly beautiful expressions for the expected cost in multi-dimensional versions of the random assignment problem. However, we do not even know how the cost scales with n.

Some of the characteristics and applications of these problems are discussed in a recent book by Burkard, Dell'Amico and Martello [BDM09]. Very little is known about the probabilistic behavior of the minimum $Z^P_{d,n}$ of $C(T)$ for $d \geq 3$, and even less is known about polynomial time algorithms for constructing good solutions. Grundel, Oliveira, Pasiliao and Pardalos [GOPP05] show that $Z^P_{d,n} \to 0$ **whp** in this case. Statistical physicists have conjectures based on "cavity" calculations [MMR04, MMR05], but there is no such nice constant as $\zeta(2)$, and no certainty even that the conjectured scaling is correct.

2 Summary of results, methods, and limitations

2.1 Planar assignment

For the Planar d-dimensional assignment problem, there is an easy lower bound of $\Omega(1/n^{d-2})$ on the expected cost (see Theorem 1). Our main result (Theorem 2) is for the case $d = 3$. Here we give an algorithm that for any constant $\varepsilon > 0$ runs in time polynomial in n and yields a solution

of expected cost $O(1/n^{1-\varepsilon})$, and where ε may be taken slowly to 0 for a "mildly exponential time" algorithm yielding an $n^{o(1)}$ approximation in the average case. Not only is this the first nearly tight upper bound obtained algorithmically, it is the only such bound except for one (see Theorem 1) following from a recent non-constructive result on hypergraph factors by Johansson, Kahn and Vu [JKV08].

Our algorithm may be thought of as an extension of one in [CS99] for 2-dimensional assignment. In that case, a bipartite matching was augmented by an alternating path of bounded length, with care taken to in regard to "conditioning" of the cost matrix. Here, partial assignments are augmented with a "bounded depth alternating path tree", a tree in which a newly added element displaces two previously selected elements, those two elements are replaced in a way displacing four selected elements, and so on, until all the displaced elements are replaced by elements in a non-conflicting, "unassigned" set (see Figure 2.1).

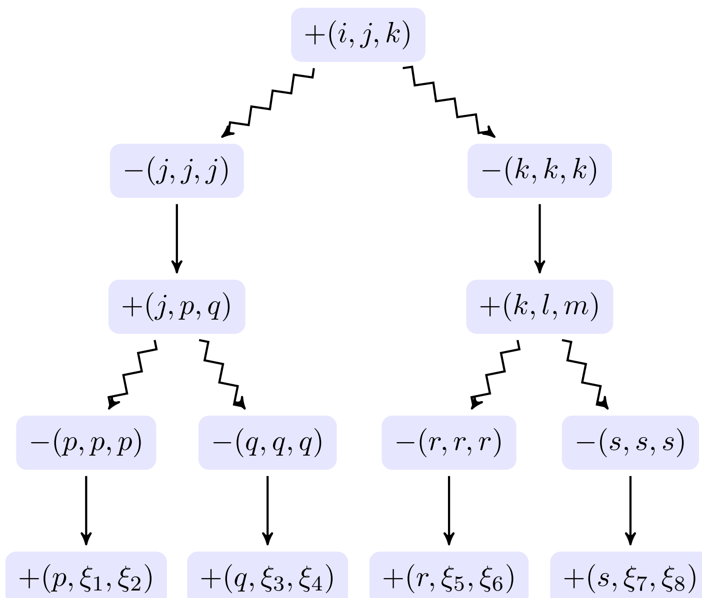

Figure 1: Diagram of alternating-path tree; see Section 3.3.2. Adding new first coordinate i to the partial assignment using hyperedge (i,j,k) implies deletion of previous assignment elements (j,j,j) and (k,k,k), with first coordinates j and k then reassigned respectively to elements (j,p,q) and (k,r,s), displacing four more existing assignment elements (p,p,p) etc., whose first coordinates are finally reassigned to unused second and third coordinates as (p,ξ_1,ξ_2) etc.

Given the tree's expansion it is surprising that this approach works, but for $d = 3$ it does. However, for $d > 3$, a more general approach fails. For any $d \geq 2$, our algorithm iterates a restricted form of local augmentation. (The main technical aspect of our result is restricting the augmentation so as to control the conditioning of the random cost matrix.) The last iteration replaces k selected assignment elements, spanning a submatrix that without loss of generality can be thought of as $[k]^d$, with $k+1$ elements spanning a sumbmatrix $[k+1]^d$ (i.e., containing the smaller submatrix). When k is a constant, the probability that there is any new such assignment of cost $\leq w$ (even without the restrictions our algorithm imposes) is at most $\binom{n}{k}(k!)^{d-1} \cdot w^k = O((nw)^k)$, so for the algorithm to have any hope of succeeding, for the last iteration alone we must budget a cost contribution $w = \Omega(n^{-1})$. That proves to be satisfactory (even taking all the iterations into account) for 2-dimensional assignment where the total cost we are aiming for is $\Theta(1)$, and for 3-dimensional assignment with its lower bound of $\Omega(n^{-1})$, but it is plainly unsatisfactory for

dimension 4, where the lower bound (which we imagine is close to the truth) is $\Omega(n^{-(d-2)}) = \Omega(n^{-2})$.

2.2 Axial assignment

Our second main result (Theorem 3) is for Axial 3-dimensional assignment, and uses a greedy algorithm: once assignment elements are chosen, they are never replaced. The 3-dimensional assignment consists of n 2-dimensional assignments, with constraints between them. (Ignoring these constraints gives a relaxed 3-dimensional problem consisting of n 2-dimensional problems. Its expected minimum cost is $n \sum_{i=1}^{n} 1/i^2$, proving that the true 3-dimensional cost has expectation $\Omega(n)$.) The first assignment is chosen greedily (at expected cost $\sum_{i=1}^{n} 1/i^2 \le \pi^2/6$) from a complete bipartite graph $K_{n,n}$. The second assignment is similar, but on a graph $K_{n,n}$ from which the edges of the first matching have been subtracted, so that each vertex has degree $n-2$, and similarly for the subsequent matchings, down to the last one where each vertex has degree 1 and the matching is forced.

The structure of the ith intermediate graph — $K_{n,n}$ minus a union of $i-1$ edge-disjoint perfect matchings — is complex, but a general result of Dyer, Frieze, and McDiarmid [DFM86] is insensitive to the details and gives an upper bound of $2n/(n-i+1)$ for the expectation of its minimum-cost perfect matching, for a total expectation of $O(n \log n)$, close to the $\Omega(n)$ lower bound. Algorithmically there is no difficulty since we are just solving a set of minimum-cost assignment problems.

As in the Planar case, though, our approach to the Axial problem falters for dimensions $d \ge 4$. Here, the 3-dimensional assignment problems are not necessarily regular, and by the end, there can even be 3-dimensional instances with no assignment at all: the greedy algorithm can fail. Perhaps this is not a problem in the average case, but at the least it necessitates a more careful analysis.

2.3 Structure of the paper

We deal with Planar assignment in Section 3, starting with easily proved non-constructive bounds (but where the upper bound relies on some heavy machinery). The next task is to analyze a "bounded depth alternating path tree algorithm" (`BDAPTA`) to prove Theorem 2. We analyze a three level version in Section 3.3, providing intuition. The general case is analyzed in Section 3.4, completing the proof.

The Axial problem is considered in Section 4. The lower bound of Theorem 3 is proved in Section 4.2, and its upper bound in Section 4.3.

3 Multi-Dimensional Planar Version

3.1 Simple bounds

Theorem 1

$$\Omega\left(\frac{1}{n^{d-2}}\right) \le Z^P_{d,n} \le O\left(\frac{\log n}{n^{d-2}}\right) \quad \textbf{whp}.$$

Proof. Clearly

$$Z^P_{d,n} \ge \sum_{i_1=1}^{n} \min_{i_2,\ldots,i_d} C_{i_1,\ldots,i_d}.$$

Each term in the above sum is distributed as $\mathrm{Exp}(n^{d-1})$ and so has expectation $1/n^{d-1}$ and variance $1/n^{2d-2}$. The Chebyshev inequality implies that the sum is concentrated around the mean.

For the upper bound we use a recent result of Johansson, Kahn and Vu [JKV08]. This implies that **whp** there is a solution that only uses d-tuples of weight at most $\frac{K \log n}{n^{d-1}}$. The upper bound follows immediately. It should be noted that their proof is non-constructive. $\square$

3.2 Main theorem

We fix the dimension to $d = 3$ for the rest of Section 3.

Theorem 2 *Suppose that $1 \leq k \leq \gamma \log_2 \log n$ where γ is any constant strictly less than 1/2. Then,* **whp***:*

(a) *Algorithm BDAPTA(k) runs in time $O(n^{2^{k+2}})$.*

(b) *The cost of the set of triples T output by BDAPTA(k) satisfies $C(T) = O(2^k n^{-1+\theta_k} \log n)$, where $\theta_k = \frac{1}{2^{k+1}-1}$.*

We are not aware of any other polynomial time algorithm that will **whp** find a solution of value $O(n^{-1+\varepsilon})$, for arbitrary positive ε.

3.3 Two Level Version of `BDAPTA`

In this section we consider a two level version of the algorithm `BDAPTA`. In this way we hope that to make it easier to understand the general version that is described in Section 3.4. With reference to Theorem 2, the two-level version means taking $k = 3$, $\theta = \theta_3 = 1/7$.

The heuristic has three phases:

3.3.1 Greedy Phase

The first phase is a simple greedy procedure.

Greedy Phase

1. Let $n_1 = n - n^{1-\theta}$, $J = K = [n]$, and $T = \emptyset$.[1]

2. For $i = 1, \ldots, n_1$ do the following:

   - Let $C_{i,j,k} = \min \{ C_{i,j',k'} : \; j' \in J, k' \in K \}$;
   - Add (i, j, k) to T and remove j from J and k from K.

At the end of this procedure the triples in T provide a partial assignment. Let

$$Z_1 = \sum_{(i,j,k) \in T} C_{i,j,k}.$$

Lemma 1

$$Z_1 \leq \frac{2}{n^{1-\theta}} \qquad \textbf{whp}.$$

[1] We will often pretend that some expressions are integer. Formally, we should round up or down but it will not matter.

Proof. We observe that if $(i,j,k) \in I$ then $C_{i,j,k}$ is the minimum of $(n-i+1)^2$ independent copies of Exp(1) and is therefore distributed as $\text{Exp}((n-i+1)^2)$. Furthermore, the random variables $C_{i,j,k}, (i,j,k) \in T$ are independent. Using the facts that an $\text{Exp}\,\lambda$ random variable has mean $1/\lambda$ and variance $1/\lambda^2$,

$$\mathbf{E}(Z_1) = \sum_{i=1}^{n_1} \frac{1}{(n-i+1)^2} \leq \int_{x=1}^{n_1+1} \frac{dx}{(n-x+1)^2} \leq \frac{1}{n^{1-\theta}}.$$

Now

$$\mathbf{Var}(Z_1) = \sum_{i=1}^{n_1} \frac{1}{(n-i+1)^4} \leq \frac{3}{n^{3(1-\theta)}} = o(\mathbf{E}(Z_1)^2)$$

and the lemma follows from the Chebyshev inequality. □

3.3.2 Main Phase

The aim of this phase is to increase the size of the partial assignment defined by T to $n - O(1)$. Let $I = I(T)$ be the set of first coordinates assigned in T, i.e., $I = I(T) = \{i : \exists j, k \text{ s.t. } (i,j,k) \in T\}$. Relabeling if necessary, without loss of generality we may assume that $I = [|T|]$. This phase will be split into *rounds*. We choose a small constant $0 < \alpha \ll 1$ and let $\beta = 1 - \alpha$. The aim of a round is to reduce the size of the set of unmatched first coordinates $X(T) = [n] \setminus I(T)$ by a factor β while increasing the total cost of the matching only by an acceptably small amount. Thus we let $x_1 = n - n_1$ and $x_t = \beta^{t-1} x_1$ for $t \geq 2$. The aim of round t is to reduce $|X(T)|$ from x_t to x_{t+1}. We continue this for $t_0 = \log_{1/\beta}(x_1/L)$ rounds where L is a large positive constant. Thus at the end of the Main Phase, if successful, we will have a partial assignment of size at least $n - 2L$.

So suppose now that we are at the start of a round and that $|X(T)| = x_t$. This is true for $t = 1$. Next let $w_0 = 2n^{-12/7} \log n$ and

$$w_t = 2n^{-6/7} x_t^{-8/7} \log^{1/7} n \qquad \text{for } t \geq 1.$$

At the start of each round we will *refresh* the array C with independent exponentials, at some cost. By this we mean that we replace C by a new array C' where $C_{i,j,k} \leq C'_{i,j,k} + w_{t-1}$ and the entries of C' are i.i.d. Exp(1)random variables. More precisely, suppose that during the previous round we determined the precise values for all $C_{i,j,k} \leq w_{t-1}$ and left our state of knowledge for the other $C_{i,j,k}$ as being at least w_{t-1}. Then the memoryless property of exponentials means that

$$C'_{i,j,k} = \begin{cases} C_{i,j,k} - w_{t-1} & \text{when } C_{i,j,k} > w_{t-1} \\ \text{fresh } X_{i,j,k} \sim \text{Exp}(1) & \text{otherwise} \end{cases}$$

has the claimed property. Thus we can start a round with a fresh matrix of independent exponentials at the expense of adding another w_{t-1} to each cost. We note also that we can **whp** carry out the Greedy Phase only looking at those $C_{i,j,k}$ of value less than w_0.

Let T_t denote the value of T at the start of round t and let $I_t = I(T_t), X_t = X(I_t)$. In round t we will add $A_t = [n - x_t + 1, n - x_{t+1}]$ to I_t. By relabeling if necessary we will assume that at the start of round t we have $T = \{(i,i,i) : 1 \leq i \leq n - x_t\}$. To add $i \in A_t$ to I_t we find distinct indices $j,k,p,q,r,s \in I_t$ (distinctness is not strictly necessary) and replace 6 of the triples in I_t by 7 new triples (see Figure 2.1:

$$+(i,j,k) - (j,j,j) - (k,k,k) + (j,p,q) + (k,r,s) - (p,p,p) - (q,q,q) - (r,r,r) - (s,s,s) + (p,\xi_1,\xi_2) + (q,\xi_3,\xi_4) + (r,\xi_5,\xi_6) + (s,\xi_7,\xi_8), \quad (1)$$

where $\xi_1,\dots,\xi_8$ are distinct members of X_t, and each of the triples added in (1) is required to have (refreshed) cost at most w_t. Roughly, we are assigning a new 1-coordinate i, this collides with previously used 2-coordinate j and 3-coordinate k, so the (j,j,j) and (k,k,k) elements are removed from the existing assignment, 1-coordinates j and k are re-added as (j,p,q) and (k,r,s) thus colliding with the previous assignment elements (p,p,p), (q,q,q), (r,r,r), and (s,s,s), and finally 1-coordinates p, q, r, s are re-added as (p,ξ_1,ξ_2) etc., where the ξ_i are elements *not* previously assigned. One may think of (1) as a binary tree version of an alternating-path construction; we will control the cost despite the tree's expansion.

Putting $W_t = w_0 + w_1 + \cdots + w_t$ we see that if we can add one element to T at a cost of at most w_t in refreshed costs, then in reality it costs us at most W_t; step (1) increases the cost by $\leq 7W_t$. Success in a round means doing this $x_t - x_{t+1}$ times, in which case the additional cost of the Main Phase will be at most 7 times

$$\sum_{t=1}^{t_0}(x_t - x_{t+1})W_t \leq x_1(w_0+w_1) + \sum_{t=2}^{t_0} x_t w_t$$
$$\leq 3n^{-6/7}\log n + 2x_1^{-1/7} n^{-6/7}\log^{1/7} n \sum_{t=2}^{t_0}\beta^{-t/7} \leq 4n^{-6/7}\log n. \quad (2)$$

We must now show that **whp** it is possible to add $x_t - x_{t+1} = \alpha x_t$ triples in round t with a (refreshed) cost of at most $7w_t$ per triple. For this we fix t and drop the suffix t from all quantities that use it. We will treat refreshed costs as actual costs and drop the word "refreshed".

We start by estimating the number of choices for assigning p. Ignoring other indices, the number of choices is distributed as the binomial $\mathrm{Bin}(\nu, 1-e^{-wx^2}) = \mathrm{Bin}(\nu,(1-o(1))wx^2)$ where $\nu = n - x$. Here $1-e^{-wx^2}$ is the probability that for a given p, there exist ξ_1,ξ_2 such that $C_{p,\xi_1,\xi_2} \leq w$. Note that

$$wx^2 = 2(x/n)^{6/7}\log^{1/7} n = o(1) \text{ and that } wnx^2 \gg \log n$$

and so the Chernoff bounds imply that, **qs**,[2] we can choose a set P of size exactly $wnx^2/2 = o(n)$, such that for each $p \in P$ there is at least one choice $\xi_1,\xi_2 \in X$ such that the triple (p,ξ_1,ξ_2) is *good*, i.e., $C_{p,\xi_1,\xi_2} \leq w$. Given this set of choices P we find that the number of choices for $q \notin P$ is distributed as the binomial $\mathrm{Bin}(\nu - |P|, 1-e^{-wx^2})$ and we can once again **qs** choose a set Q, disjoint from P such that $|Q| = wnx^2/2$ and each $q \in Q$ is in some good triple (q,ξ_3,ξ_4) where $\xi_3,\xi_4 \in X$. Similarly, we can choose sets R, S of choices for r, s, of size $wnx^2/2$, such that P,Q,R,S are pairwise disjoint.

Observation 2 *Each $\xi \in X$ is in $\mathrm{Bin}(x\nu, 1-e^{-w})$ good triples of the form $(p \in P, \xi', \xi'')$ and so* **qs** *it is in at most*

$$2wnx = \frac{4n^{1/7}\log^{1/7} n}{x^{1/7}}$$

such triples.

We now discuss our choices for j and k. For a fixed j there are $w^2n^2x^4/4$ pairs in $P \times Q$ and each has a probability $1-e^{-w}$ of forming a good triple (j,p,q). Let j be *useful* if there is such a pair and *useless* otherwise. Then

$$\mathbf{Pr}(j \text{ is useless}) \leq \exp\left\{-\frac{w^3n^2x^4}{4}\right\} \leq 1 - \frac{w^3n^2x^4}{5}.$$

[2]A sequence of events $\mathcal{E}_n, n \geq 0$ are said to occur *quite surely*, **qs**, if $\mathbf{Pr}(\mathcal{E}_n) = 1 - O(n^{-K})$ for any constant $K > 0$.

It follows that the number of useful $j \notin Y = P \cup Q \cup R \cup S$ dominates $\mathrm{Bin}(n - o(n), w^3 n^2 x^4/5)$ and so **qs** we can choose a set J of useful $j \notin Y$ of size

$$\frac{w^3 n^3 x^4}{6} = \frac{4n^{3/7} x^{4/7} \log^{3/7} n}{3} = o(n).$$

We can by a similar argument choose a set K of useful k of this size disjoint from J and Y.

Observation 3 *A fixed p is in at most $\mathrm{Bin}(wn^2x^2/2, 1 - e^{-w})$ good triples (j, p, q) where $(j, q) \in J \times Q$ and so* **qs** *every p is in at most $w^2 n^2 x^2$ such triples.*

Suppose then that in the middle of a round we have added $y < \alpha x$ triples to T. The number of $\xi \in X$ that can be used in a good triple (p, ξ, η) will have been reduced by y. The number of η will have been reduced by the same amount. It follows from Observation 2 that the number of choices for p will have been reduced by at most $2\alpha x \times 2wnx$. By Observation 3 this reduces the number of choices for j by at most $2\alpha x \times 2wnx \times w^2 n^2 x^2 + 7\alpha x \ll |J| = w^3 n^3 x^4/6$. The additional term $+7\alpha x$ accounts for the choices we lost because they have previously been used in this round. So our next i will get a choice of at least $\mathrm{Bin}((w^3 n^3 x^4/7)^2, 1 - e^{-w})$ choices for a good triple (i, j, k). So the expected number of choices is at least $w^7 n^6 x^8/49 = (2^7/49)\log n$ and then the probability there is no choice is $o(n^{-1})$. This is sufficient to ensure that **whp** there is always at least one choice for every i.

3.3.3 Final Phase

We now have to add only $O(1)$ indices to I. At this point there is a problem with the bottom-up approach of the previous phase if $x < 8$, clearest in the case $x = 1$, say the single element n, when each of $\xi_1, \ldots, \xi_8$ would have to be n, leading to an illegal assignment. Thus instead we will work top down. The details of this will cause more conditioning of the matrix, and therefore we refresh C after each increase in I, at an extra cost of $w = Kn^{-6/7} \log^{1/7} n$. So, if successful, the cost of this round is $O(W_{t_0} + w) = O(n^{-6/7} \log^{1/7} n)$.

Let us now replace the notation of (1) by

$$+(i,j,k)-(j_1,j,j_3)-(k_1,k_2,k)+(j_1,p,q)+(k_1,r,s)-(p_1,p,p_3)-(q_1,q_2,q)-(r_1,r,r_3)-(s_1,s_2,s)+ (p_1,i_2,p_3)+(q_1,q_2,j_3)+(r_1,s_2,i_3)+(s_1,k_2,r_3), \quad (3)$$

where any subtracted triple such as (j_1, j, j_3) denotes a previous match (we are no longer assuming the convention that such a triple would be (j, j, j)), and where i_2, i_3 are unused 2- and 3-coordinates respectively.

Fix j (and thus its previously matched companion indices j_1, j_3) and let Z_j be the number of choices for p, q (with their previously matched companion indices p_1, p_3, q_1, q_2) such that $C(j_1, p, q), C(p_1, i_2, p_3), C(q_1, q_2, j_3) \leq w$. This has the distribution $B_1(B_2(n, w)B_3(n, w), w)$ where B_1, B_2, B_3 denote independent binomials, with B_2 counting the good choices for p, B_3 those for q, and B_1 those for j using these p and q possibilities. Using Chernoff bounds on the binomials B_2, B_3 we see that **whp** Z_j dominates $B(n^2w^2/2, w)$ which dominates $\mathrm{Be}(n^2w^3/3)$, the Bernoulli random variable that is 1 with probability $n^2w^3/3$ and 0 otherwise. The same holds for index k and (3) has been constructed so that choices for j, k are independent. So, the number of choices for j, k dominates $\mathrm{Bin}(n^2, w(n^2w^3/3)^2)$ which has expectation $\Omega(\log n)$ and so is non-zero **whp**.

This completes the analysis of `BDAPTA` when there are two levels.

3.4 General 3-Dimensional Version

We follow the same three phase strategy. k is a positive integer, $2 \le k \le \gamma \log\log n$.

3.4.1 Greedy Phase

This is much as before. Proceed as in Section 3.3.1 but taking $\theta = \theta_k$ (recall θ's definition from Theorem 2) and defining n_1 accordingly. Lemma 1 continues to hold.

3.4.2 Main Phase

Let

$$\alpha = 2^{-2k-2}\left(1 - \sqrt{2/3}\right)$$

and let β, t_0 and $x_t, t = 1, \ldots, t_0$ be defined as in Section 3.3.2. Let I_t, X_t, A_t have the same meaning as well. Now let $w_0 = 2n^{-2(1-\theta_k)} \log n$ and

$$w_t = 2x_t^{-1-\theta_k} n^{\theta_k - 1} \log^{\theta_k} n \qquad \text{for } t \ge 1$$

and

$$W_t = w_0 + w_1 + \cdots + w_t = O\left(\frac{\log^{\theta_k} n}{n^{1-\theta_k}}\right).$$

The aim of round t is once again to add $x_t - x_{t+1}$ new indices to I_t using triples with (refreshed) cost at most w_t. We will assume that at the start of round t we have $T = \{(i,i,i) : \ 1 \le i \le n - x_t\}$. In analogy with (1), to add $i \in A_t$ to I_t we will add $2^{k+1} - 1$ triples to T and remove $2^{k+1} - 2$ triples, in which case the additional cost of the Main Phase will be at most $2^{k+1} - 1$ times

$$\sum_{t=1}^{t_0} (x_t - x_{t+1}) W_t \le x_1(w_0 + w_1) + \sum_{t=2}^{t_0} x_t w_t$$
$$\le 3n^{\theta_k - 1} \log n + 2x_1^{-\theta_k} n^{\theta_k - 1} \log^{\theta_k} n \sum_{t=2}^{t_0} \beta^{-\theta_k t} \le 4n^{\theta_k - 1} \log n. \quad (4)$$

The notation used in (1) is obviously insufficient. We imagine a rooted tree Γ of triples. The root will be $\rho = (i_0, j_0, k_0)$ where i_0 is the index to be added to I_t. The root is at level zero. The triples at odd levels are to be deleted from T and the vertices at even levels are to be added to T. Every triple at an odd level $2l - 1$ will therefore have the form (p,p,p) where $p \in I_t$. This triple will have one child (p,a,b) which will replace the parent triple in 1-plane p. If $l < k$ then $a, b \in I_t$ and if $l = k$ then $a, b \in X_t$. A triple $u = (p,a,b)$ at an even level will have two children. By construction, u will be the unique triple in 1-plane p, but now we will have two triples in 2-plane a and 3-plane b. Thus the children of u are (a,a,a) and (b,b,b). This defines a tree corresponding to adding $2^{k+1} - 1$ and removing $2^{k+1} - 2$ triples from T. We ensure that if $u = (p,a,b)$ is a triple at an even level, then p, a, b do not appear anywhere else in the tree, except at the child of u as previously described. We do this so that additions in one part of the tree do not clash with additions in another part and then the additions and deletions give rise to a partial assignment. We also insist that if $u = (p,a,b)$ is a triple at an even level then $C_{p,a,b} \le w$. We call such a tree *feasible*. We considered each level of Γ to be ordered so it makes sense to talk of the rth vertex of level $2l$ where $1 \le r \le 2^l$.

We now have to show that **whp** there is always at least one such tree Γ for each $i \in A_t$. We take the same *bottom-up* approach that we did in Section 3.3. We fix t and drop the suffix t from

all quantities that use it. We start by estimating the number of choices for a p that can be in a triple (p, x, y) at level $2k$. Ignoring other indices, the number of choices is again distributed as the binomial $\mathrm{Bin}(\nu, 1 - e^{-wx^2}) = \mathrm{Bin}(\nu, (1 - o(1))wx^2)$ where $\nu = n - x = n - o(n)$. Note that $wx^2 = K(x/n)^{1-\theta_k} \log^{\theta_k} n = o(1)$ and that $wnx^2 = \tilde{\Omega}(n^{\theta_k}) \gg \log n$. (Here our notation $f(n) \gg g(n)$ means that $f(n)/g(n) \to \infty$ with n). So the Chernoff bounds imply that **qs** we can choose a set P of size exactly $wnx^2/2 = o(n)$, such that for each $p \in P$ there is at least one choice ξ_1, ξ_2 such that the triple (p, ξ_1, ξ_2) is *good*, i.e., $C_{p,\xi_1,\xi_2} \leq w$. We will in fact be able to choose 2^k disjoint sets $P_{l,k}, 1 \leq l \leq 2^k$ since replacing ν by $\nu - 2^k wnx^2/2$ will not significantly change the above calculations. (Here $2^k wnx^2 = O(n^{1-\theta_k+\theta_k^2} \log^{\theta_k+\gamma} n) = o(n)$).

Observation 4 *Each $\xi \in X$ is in $\mathrm{Bin}(x\nu, 1 - e^{-w})$ good triples of the form $(p \in P_{l,k}, \xi, .)$ and so* **qs** *it is in at most $2wnx$ such triples. (Here $wnx = 2\left(\frac{n \log n}{x}\right)^{\theta_k} \gg \log n$).*

Let

$$\nu_0 = wnx^2/2 \text{ and } \nu_{l+1} = wn\nu_l^2/2 \text{ for } 0 \leq l < k. \tag{5}$$

The solution to this recurrence is

$$\nu_l = \left(\frac{wn}{2}\right)^{2^{l+1}-1} x^{2^{l+1}} = (n \log n)^{(2^{l+1}-1)\theta_k} x^{(2^{k+1}-2^{l+1})\theta_k}.$$

Observe that ν_l increases with l. Note also that if $l \leq k - 2$ then

$$w\nu_l^2 \leq w\nu_{k-2}^2 = 2\left(\frac{x}{n}\right)^{2^k\theta_k} \log^{(2^k-1)\theta_k} n = o(1), \tag{6}$$

$$wn\nu_l \geq wn\nu_0 = \frac{w^2n^2x^2}{2} = 2\left(\frac{n \log n}{x}\right)^{2\theta_k} \gg \log n. \tag{7}$$

We now have the basis for an inductive claim that **qs** if $l \leq k - 1$ and $u = (p, a, b)$ is a triple at an even level $2(k - l)$ then there are at least ν_l choices for p such that there exists a triple $u = (p, a, b)$ with $C_u \leq w$ and a feasible tree Γ_u with u as root and depth $2l + 1$. Our analysis above has proved the base case of $l = 0$. Imagine now that we are filling in the possibilities for the rth triple (p, a, b) at level $k - l$. We fill in these possibilities level by level starting at level $2k$. Imagine also that we have identified ν_{l-1} choices for each of a, b. This can be an inductive assumption, so for example a will have to be a possible selection for the first component of the $(2r - 1)$st triple at level $2(k - (l - 1))$.

For a fixed p, conditional on our having selected exactly ν_{l-1} choices A, B for a, b, let p be *useful* if there is a pair $(a, b) \in A \times B$ with $C_{p,a,b} \leq w$ and *useless* otherwise. Then, using (6),

$$\mathbf{Pr}(p \text{ is useless}) \leq \exp\left\{-w\nu_{l-1}^2\right\} \leq 1 - \frac{2w\nu_{l-1}^2}{3}.$$

It follows that the number of useful p that have not been previously selected dominates $\mathrm{Bin}(n - o(n), 2w\nu_{l-1}^2/3)$. Here $o(n) = \sum_{s \leq l} 2^{k-s} w\nu_s^2$ bound the number of *forbidden* p's. It follows that **qs** we can choose a set of useful p's of size $\nu_l = wn\nu_{l-1}^2/2 = o(n)$. We can do this so that each node of Γ gets distinct choices.

Observation 5 *A fixed a is in at most $\mathrm{Bin}(n\nu_{l-1}/2, 1 - e^{-w})$ good triples (p, a, b) feasible for level $2(k - l)$ and so* **qs** *every a is in at most $wn\nu_{l-1}$ such triples, see* (7).

This completes our induction. We now apply the above to show that round t succeeds **whp**.

Suppose that in the middle of a round we have added $y < \alpha x$ triples to T. The number of $\xi \in X$ that can be used in a good triple (p, ξ, η) at level $2k$ will have been reduced by y. Thus the number of choices for p in any triple in this level will have been reduced by at most $2^k \times 2 \times \alpha x \times 2wnx$, see Observation 4. This reduces the number of choices for p in a triple at level $2(k-1)$ by at most $2^{k+2}\alpha wnx^2 \times wn\nu_0 = 2^{k+3}\alpha wn\nu_0^2$, see Observation 5. So let μ_l denote the number of choices for p in triples $p(,.,.)$ at level $2(k-l)$ that are forbidden by choices further down the tree. We have just argued that $\mu_1 \leq 2^{k+3}\alpha wn\nu_0^2$. In general we can use Observation 5 to conservatively argue that

$$\mu_l \leq wn\nu_{l-1}(\mu_{l-1} + 2^{k+1}\alpha x).$$

It follows that for $l \geq 2$ we have

$$\frac{\mu_l}{\nu_l} \leq 2\frac{\mu_{l-1}}{\nu_{l-1}} + \frac{2^{k+2}\alpha x}{\nu_{l-1}} \leq 2\frac{\mu_{l-1}}{\nu_{l-1}} + 2^{k+2}\alpha\left(\frac{x}{n\log n}\right)^{(2^l-1)\theta_k} \leq 2\frac{\mu_{l-1}}{\nu_{l-1}} + 2^{k+2}\alpha\left(\frac{x}{n\log n}\right)^{\theta_k}.$$

It follows that

$$\frac{\mu_{k-1}}{\nu_{k-1}} \leq 2^{k-2}\frac{\mu_1}{\nu_1} + 2^{2k+1}\alpha\left(\frac{x}{n\log n}\right)^{\theta_k} \leq 2^{2k+2}\alpha.$$

We see that at the root there will still be at least $(1-2^{2k+2}\alpha)\nu_{k-1}$ choices for j_0, k_0. So i_0 will get a choice of at least $\mathrm{Bin}((1-2^{2k+2}\alpha)^2\nu_{k-1}^2, 1-e^{-w})$ choices for a good triple (i_0, j_0, k_0). So the expected number of choices is at least $2w\nu_{k-1}^2/3$, our choice of α implies this. Now $w\nu_{k-1}^2 = 2\log n$ and this is sufficient to ensure that **whp** there is always at least one choice for every i_0.

3.4.3 Final Phase

We can execute the Main Phase so long as $x \geq 2^k$. Now assume that $1 \leq x < 2^k$. We now have to add only $O(1)$ indices to I. This time we refresh C an $O(2^k)$ number of times at an extra cost of $w_f = \dfrac{\log^{\theta_k} n}{n^{1-\theta_k}}$ each time we add an index. So, if successful, the cost of this round is $O(W_{t_0} + w_f) = O\left(\frac{\log^{\theta_k} n}{n^{1-\theta_k}}\right)$.

We first make an inductive assumption: We have a partial assignment I where $|I| \leq n-2$. (The reader might think that we should assume $|I| \leq n-1$, but here we use the induction hypothesis after one more index has temporarily been deleted from I, prior to a replacement). Assume that the matrix C is unconditioned and $i \notin I$: Then we can in $O(n^{2\ell})$ time **whp** find a set P of size $\nu_{\ell-1}$ (with $x = 1$ in definition (5)) and a collection $Q_p, p \in P$ of sets of size $\nu_{\ell-1}$ such that for each $(p, q \in P_p)$ there is an assignment P' with $(i,p,q) \in P'$ and $I(P') \supsetneq I(P)$ and $C(P') = C(P) + C(i,p,q) + O(w_f)$. This is true for $\ell = 1$ since we can make the changes

$$+(i,p,q) - (p_1,p,p_3) - (q_1,q_2,q) + (p_1,i_2,p_3) + (q_1,q_2,i_3)$$

where i_2, i_3 are unused 2- and 3-coordinates respectively. The number of choices for p, q are independent $\mathrm{Bin}(n, w)$.

For the inductive step, we first refresh the matrix C. Then for each $p \in [n]$ we let $I' = I - \{p\}$ and apply the induction hypothesis to generate $\nu_{\ell-2}^2$ choices of assignment that add back p_1 to I'. We find that **whp** at least $w\nu_{\ell-2}^2/2 = \nu_{\ell-1}$ of these have $C(p_1,.,.) \leq w$. Let this set be P. Now refresh C again and apply the same argument for each $p \in P$ to generate choices Q_p for p. This completes the induction.

Now let $\ell = k$ and refresh C one more time. Let $P, Q_p, p \in P$ be the sets of size ν_{k-1} promised by the above argument. We have $\mathrm{Bin}(\nu_{k-1}^2, w)$ choices of j, k which can be used to add $i \notin I$ to I at a cost of $O(w)$. In expectation this is $2 \log n$ and so we succeed **whp**.

For the execution time of the algorithm we observe that to add a triple to I involves replacing $2^{k+1} - 2$ triples by $2^{k+1} - 1$ triples. This involves the construction of a tree T where for the $2^{k+1} - 1$ vertices that we add, we have n^2 choices to make, in the worst-case and for which the other choices are forced. This gives a running time of $O\left(n \times (n^2)^{2^{k+1}-1}\right)$ time.

This completes the proof of Theorem 2. □

4 Multi-Dimensional Axial Version

4.1 Main theorem

Here we give our main theorem for Axial assignment.

Theorem 3 *The optimal solution value $Z_{d,n}^A$ satisfies the following:*

(a) $Z_{d,n}^A = \Omega(n^{d-2})$ **whp** *for $d \geq 3$.*

(b) *When $d = 3$ there is a polynomial time algorithm that finds a solution with cost Z where $Z = O(n \log n)$* **whp**.

The theorem is proved in the next two sections.

4.2 Lower bound

It is clear that $Z_{d,n}^A \geq Z_1 + Z_2 + \cdots + Z_{n^{d-2}}$ where Z_i is the minimum cost of the 2-dimensional assignment with cost matrix $A_{j,k} = C_{i_1,\ldots,i_{d-2},j,k}$. We know that $Z_j \geq (1 - o(1))\zeta(2)$ **whp** and the Z_i's are independent. It follows that **whp** $Z_{3,n}^A \geq (1 - o(1))n^{d-2}\zeta(2) > 3n^{d-2}/2$.

4.3 Upper bound for $d = 3$

For the upper bound we need a result of Dyer, Frieze and McDiarmid [DFM86]. We will not state it in full generality, instead we will tailor its statement to precisely what is needed. Suppose that we have a linear program

$$P: \qquad \text{Minimize } c^T x \text{ subject to } Ax = b,\ x \geq 0.$$

Here A is an $m \times n$ matrix and the cost vector $c = (c_1, c_2, \ldots, c_n)$ is a sequence of independent copies of $\mathrm{Exp}(1)$. Let Z_P denote the minimum of this linear program. Note that Z_P is a random variable. Next let y be *any* feasible solution to P.

Theorem 4 ([DFM86])

$$\mathbf{E}(Z_P) \leq m \max_{j=1,2,\ldots,n} y_j. \tag{8}$$

Furthermore, Z_P is at most $1 + o(1)$ times the RHS of (8), **whp**.

Now consider the following greedy-type algorithm. We find a minimum 2-dimensional assignment for 1-plane $i = 1$, we then find a minimum assignment for 1-plane $i = 2$, consistent with choice for 1-plane $i = 1$, and so on:

Greedy

1. For $i = 1, \ldots, n$ do the following:
   - Let $G = K_{n,n} \setminus (M_1 \cup M_2 \cup \cdots M_{i-1})$;
   - If $(j,k) \in E(G)$ let $A_{j,k} = C_{i,j,k}$.
   - Let M_i be a minimum cost matching of G using edge weights A.

The output, $M_1, M_2, \ldots, M_n$ defines a set of triples $T = \{(i,j,k) : (j,k) \in M_i\}$. We claim that if $Z_i = A(M_i)$ then

$$\mathbf{E}(Z_i) \leq \frac{2n}{n-i+1}. \tag{9}$$

For this we apply Theorem 4 to the linear program

$$\begin{aligned} \text{Minimize} \sum_{(j,k)\in E(G)} A_{j,k} x_{j,k} \quad & \text{subject to} \\ \sum_{k:\,(j,k)\in E(G)} x_{j,k} = 1, \quad & j = 1,2,\ldots,n \\ \sum_{j:\,(j,k)\in E(G)} x_{j,k} = 1, \quad & k = 1,2,\ldots,n \\ x_{j,k} \geq 0, \quad & j,k = 1,2,\ldots,n. \end{aligned}$$

We note that there are $2n$ constraints and that $x_{j,k} = 1/(n-i+1)$ is a feasible solution. With Theorem 4, this implies (9) and the upper bound in Theorem 3 for the case $d = 3$.

5 Conclusions

For the 2-dimensional random assignment problem, we know the limiting expected cost, and a given instance can be solved in polynomial time. As noted in the Introduction, much less is known about multidimensional assignment problems, and as far as we are aware, nothing was known about polynomial-time algorithms solving these problems well on average. For the 3-dimensional Axial assignment problem, the present paper is the first to prove an upper bound within a logarithmic factor of the obvious $O(n)$ lower bound (likely the true answer), doing so by analyzing a simple and fast greedy algorithm. For the 3-dimensional Planar assignment problem, we give the first upper bounds within polylogarithmic factors of the obvious $O(1/n)$ lower bound. One of our upper bounds is a trivial application of a result of Johansson, Kahn and Vu [JKV08]; we cannot really take credit for it. The second, however, comes from analyzing an algorithm that is the first that solves this problem well, on average. For any $\varepsilon \geq (\log n)^{-\gamma}$, for any constant $\gamma > 1/2$, run for "mildly exponential" time (polynomial time if ε is constant), the algorithm delivers a solution within a factor $O(n^{\varepsilon})$ of the expected minimum; we know of no other such approximation. As discussed in Section 2, our results do not extend to $d \geq 4$.

We are left with open questions including these:

P1 What are the growth rates of $\mathbf{E}[Z^P_{d,n}]$ and $\mathbf{E}[Z^A_{d,n}]$ for $d \geq 3$?

P2 Are there asymptotically optimal, polynomial time algorithms for solving these problems when $d \geq 3$?

P3 For $d > 3$, are there polynomial time algorithms yielding solutions within logarithmic or $O(n^e)$ factors for Axial and Planar assignment problems (as we have given for $d = 3$)?

P4 Frieze [Fri74] gave a bilinear programming formulation of the 3-dimensional planar problem. There is a natural heuristic associated with this formulation (see Appendix). What are its asymptotic properties?

References

[Ald92] David Aldous, *Asymptotics in the random assignment problem*, Probab. Theory Related Fields **93** (1992), no. 4, 507–534.

[Ald01] David J. Aldous, *The* $\zeta(2)$ *limit in the random assignment problem*, Random Structures Algorithms **18** (2001), no. 4, 381–418.

[BDM09] Rainer Burkard, Mauro Dell'Amico, and Silvano Martello, *Assignment problems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2009.

[CS99] Don Coppersmith and Gregory B. Sorkin, *Constructive bounds and exact expectations for the random assignment problem*, Random Structures Algorithms **15** (1999), no. 2, 113–144.

[DFM86] Martin E. Dyer, Alan M. Frieze, and Colin J.H. McDiarmid, *On linear programs with random costs*, Mathematical Programming **35** (1986), 3–16.

[Don69] W.E. Donath, *Algorithm and average-value bounds for assignment problems*, IBM Journal of Research and Development **13** (1969), no. 4, 380–386.

[Fri74] A. M. Frieze, *A bilinear programming formulation of the* 3*-dimensional assignment problem*, Math. Programming **7** (1974), 376–379.

[Fri83] ———, *Complexity of a* 3*-dimensional assignment problem*, European J. Oper. Res. **13** (1983), no. 2, 161–164.

[Fri85] ———, *On the value of a random minimum spanning tree problem*, Discrete Appl. Math. **10** (1985), no. 1, 47–56.

[FY81] A.M. Frieze and J. Yadegar, *An algorithm for solving 3-dimensional assignment problems with application to scheduling a teaching practice*, Journal of the Operational Research Society (1981), no. 32, 989–995.

[GOPP05] Don Grundel, Carlos A. S. Oliveira, Panos M. Pardalos, and Eduardo Pasiliao, *Asymptotic results for random multidimensional assignment problems*, Comput. Optim. Appl. **31** (2005), no. 3, 275–293. MR MR2160697

[JKV08] Anders Johansson, Jeff Kahn, and Van Vu, *Factors in random graphs*, Random Structures Algorithms **33** (2008), no. 1, 1–28.

[Kar72] Richard M. Karp, *Reducibility among combinatorial problems*, Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972), Plenum, New York, 1972, pp. 85–103.

[LW04] Svante Linusson and Johan Wästlund, *A proof of Parisi's conjecture on the random assignment problem*, Probab. Theory Related Fields **128** (2004), no. 3, 419–440.

[MMR04] O. C. Martin, M. Mezard, and O. Rivoire, *A frozen glass phase in the multi-index matching problem*.

[MMR05] O. C. Martin, M. Mézard, and O. Rivoire, *Random multi-index matching problems*, J. Stat. Mech. Theory Exp. (2005), no. 9, P09006, 36 pp. (electronic).

[MP85] M. Mézard and G. Parisi, *Replicas and optimization*, J. Physique Lettres **46** (1985), 771–778.

[MP87] ______, *On the solution of the random link matching problems*, J. Physique Lettres **48** (1987), 1451–1459.

[NPS05] Chandra Nair, Balaji Prabhakar, and Mayank Sharma, *Proofs of the Parisi and Coppersmith-Sorkin random assignment conjectures*, Random Structures Algorithms **27** (2005), no. 4, 413–444. MR MR2178256 (2006e:90050)

[Par98] Giorgio Parisi, *A conjecture on random bipartite matching*, Physics e-Print archive, http://xxx.lanl.gov/ps/cond-mat/9801176, January 1998.

[Wäs09] Johan Wästlund, *An easy proof of the zeta(2) limit in the random assignment problem*, Electronic Communications in Probability (2009), no. 14, 261–269.

A Bilinear Programming Formulation

Frieze [Fri74] re-formulated the 3-dimensional planar problem as

$$\text{Minimize} \sum_{i,j,k=1}^{n} C_{i,j,k} y_{i,j} z_{i,k} \text{ subject to } x, y \in P_A$$

where P_A is the bipartite matching polyhedron $\sum_{i=1}^{n} x_{i,j} = 1 = \sum_{j=1}^{n} x_{i,j}$, for all $1 \leq i, j \leq n$.

Now denote the objective above by $C(y, z)$. The following heuristic was used successfully in a practical situation [FY81]:

1. Choose y_0, z_0 arbitrarily; $Z_0 = C(y_0, z_0)$; $i = 0$.
2. Repeat until $Z_{i+1} = Z_i$.
    - Let y_{i+1} maximize $C(y, z_i)$.
    - Let z_{i+1} maximize $C(y_{i+1}, z)$.
    - $Z_{i+1} = C(y_{i+1}, z_{i+1})$.
    - $i = i + 1$.